<table>
<thead>
<tr>
<th>Title</th>
<th>Fuzzy EMG classification for prosthesis control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Chan, FHY; Yang, YS; Lam, FK; Zhang, YT; Parker, PA</td>
</tr>
<tr>
<td>Citation</td>
<td>IEEE Transactions on Neural Systems and Rehabilitation Engineering, 2000, v. 8 n. 3, p. 305-311</td>
</tr>
<tr>
<td>Issued Date</td>
<td>2000</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10722/42858">http://hdl.handle.net/10722/42858</a></td>
</tr>
<tr>
<td>Rights</td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; ©2000 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
</tbody>
</table>
Fuzzy EMG Classification for Prosthesis Control

Francis H. Y. Chan, Yong-Sheng Yang, F. K. Lam, Yuan-Ting Zhang, and Philip A. Parker

Abstract—This paper proposes a fuzzy approach to classify single-site electromyograph (EMG) signals for multifunctional prosthesis control. While the classification problem is the focus of this paper, the ultimate goal is to improve myoelectric system control performance, and classification is an essential step in the control. Time segmented features are fed to a fuzzy system for training and classification. In order to obtain acceptable training speed and realistic fuzzy system structure, these features are clustered without supervision using the Basic Isodata algorithm at the beginning of the training phase, and the clustering results are used in initializing the fuzzy system parameters. Afterwards, fuzzy rules in the system are trained with the back-propagation algorithm. The fuzzy approach was compared with an artificial neural network (ANN) method on four subjects, and very similar classification results were obtained. It is superior to the latter in at least three points: slightly higher recognition rate; insensitivity to overtraining; and consistent outputs demonstrating higher reliability. Some potential advantages of the fuzzy approach over the ANN approach are also discussed.

Index Terms—Classification, electromyography (EMG), fuzzy logic, neural network, prosthesis.

I. INTRODUCTION

ELECTROMYOGRAPHY (EMG) pattern recognition has been applied in controlling prosthetic devices for amputees [4], [6]. The control is based on the fact that amputees still have the phantom of functions and hence are able to generate a repeatable (although perhaps gradually varying) EMG pattern corresponding to each of the functions. In the past decades, much research has been done on the recognition of EMG signals [4]–[10], [13], [16], most of which has been reviewed by Hudgins, etc. [5], [6] The main prosthesis control functions of interest were flexion and extension of the forearm and pronation and supination of the wrist respectively. The hand actions such as hand opening and grasping, and finger bending were also of some interest. In order to differentiate these control functions, investigators developed various EMG features which include EMG signal amplitude, zero-crossing [5], [13], [16], EMG frequency characteristics [7]–[9], and coefficients of an EMG autoregressive model [4]. The classification tools covered linear discriminate functions [13], [16], neural networks [5], [8], [9], and fuzzy systems [7]. Researchers have applied different kinds of mathematical models and pattern recognition techniques to the problem; however, they are not yet commercially available. Table I gives a brief summary of some typical EMG recognition systems.

EMG classification systems as listed above suffer from one or more of the following drawbacks: large number of electrodes; sensitivity to electrode displacement; low recognition rate; perceivable delay in control (delay > 300 ms; a 200 to 300 ms interval is a clinically recognized maximum delay that users find acceptable before they get frustrated with the slow response of the prosthesis [17]). The system proposed by Hudgins et al. [5] has solved most of the above difficulties; however, its recognition rate is subject-dependent, ranging from 70 to 98%. We have used the same ANN method as that in [5] with data sets from the same four subjects and achieved similar classification results to those reported by Hudgins’ paper. However, we noted that the success of the method is chiefly due to the appropriately selected features.

Fuzzy logic systems are advantageous in biomedical signal processing and classification. Biomedical signals are not always strictly repeatable, and may sometimes even be contradictory. One of the most useful properties of fuzzy logic systems is that contradictions in the data can be tolerated. Furthermore, using trainable fuzzy systems, it is possible to discover patterns in data which are not easily detected by other methods, as can also be done with neural network. Finally, the experience of medical experts can be incorporated. It is possible to integrate this incomplete but valuable knowledge into the fuzzy logic system due to the system’s reasoning style, which is similar to that of a human being. This is a significant advantage over the artificial neural network (ANN). In order to search for an improved solution to the EMG classification problem, this paper proposes a new fuzzy approach for EMG recognition based on most of the same time-segmented features as used by Hudgins et al. [5]; the results are compared to those of the ANN method [5].

Hudgins et al. [5] applied electrodes consisting of a single differential channel with an active electrode over each of the biceps brachii and triceps muscles (for further detail re electrodes see [5]). They used segmented features of EMG signals
and an ANN for one-site multi-functional prosthetic control. The classification strategy is illustrated in Fig. 1. The four functions to be classified and controlled were: 1) elbow extension, 2) elbow flexion, 3) wrist pronation, and 4) wrist supination. The system’s correct classification rate was between 70 and 98% among 80 test patterns presented after an initial training of the network. The EMG signal was acquired by a single bipolar electrode pair placed at the biceps and the triceps. The moving average of the signal was monitored. When a threshold was exceeded, 200 samples (at 1 kHz sampling rate) of EMG samples were stored. The EMG signals were divided into several time segments to preserve pattern structure, and features were extracted from these segments. The features included mean absolute value (MAV), mean absolute value slope (MAVSLP), number of zero-crossings (ZC), slope sign changes (SSC), and wave length or wave complexity (WC) [5]. These time-segmented features were fed to a three-layer ANN which was able to generalize from a small training data set (typically 20 trials of each function for training). The hidden-layer contained eight nodes, as derived from the performance vs. hidden node number curve acquired via experiments. Each output in the output layer corresponded to one function. The required output was set to 0.9 for function selected and 0.1 for others in the training phase. Based on experiments, we use 0.4 as the activation threshold so that the maximum output of the network has to exceed 0.4 for the activated function. The ANN method was successfully reproduced so that our fuzzy approach can be compared to it.

Section II of this paper introduces the trainable fuzzy system, presents our fuzzy EMG classification scheme, and details why and how we have improved the fuzzy system. Section III presents experimental results with fuzzy EMG classification and comparative analysis with the ANN method. Section IV presents analysis of the fuzzy system’s superiority, training time comparison, and the overtraining problem.

II. EMG CLASSIFICATION WITH A TRAINABLE FUZZY LOGIC SYSTEM

In the last decade soft computing, mainly composed of fuzzy logic, neural networks, and genetic algorithms, has achieved great success in many applications. Unlike traditional hard computing, soft computing exploits tolerance of imprecision, uncertainty, and partial truth to achieve tractable, robust, and low-cost solutions to decision problems. The neural network is able to learn through the acquired data and to make generalizations on new in-coming data, and thus mimic human decision process to some extent. The multilayer network (MLN) has been successfully applied in solving pattern recognition problems, using the back error propagation algorithm in training. Fuzzy logic systems emulate human decision-making more closely than the ANN. The fuzzy systems at first fuzzify inputs into membership degrees of fuzzy sets, and then infer by fuzzy logic through rules which usually come from experience. Fuzzy systems have been widely applied in solving control problems. In short, soft computing approaches are good at coping with imprecisely-defined situations in a model-free way, without having to model a system beforehand. In other words, they are sufficiently “soft” to be versatile models.

The kernel of a fuzzy system is the fuzzy inference engine. The knowledge of an expert or well-classified examples are expressed as or transferred to a set of “fuzzy production rules” of the form IF-THEN, leading to algorithms describing what action or selection should be taken based on the currently observed information. In the fuzzy method, nothing is done at random. Information containing a certain amount of vagueness is expressed as faithfully as possible, without the distortion of forcing it into a “crisp” mold, and it is then processed in a proper manner.

A. The Structure and Training of Fuzzy System

The fuzzy system used in the present paper has been reported by Wang [18]. Fig. 2 shows the basic configuration of a fuzzy logic system with a fuzzifier and a defuzzifier. The fuzzy logic system with center average defuzzifier, product-inference rule, nonsingleton fuzzifier, and Gaussian membership function can be described by the following function [18]:

$$f(x) = \frac{1}{\sum_{j=1}^{M} y_j^{n}} \left( \prod_{j=1}^{n} \alpha_j \exp \left( -\frac{(x_i - \mu_i)^2}{\sigma_i^2} \right) \right)$$

This system is used in the present paper. It has been shown that the system can be a “universal approximator” in a manner similar to the ANN [18].

The fuzzy system described above can be represented by a three-layer feed-forward network (but different from the MLN) as depicted in Fig. 3. It resembles a radial basis neural network (RBN) [11]. However, both the fuzzy set definition (represented by centroid and width) and the rules (represented by the connecting weights) are trained by the back-propagation algorithm in the fuzzy system, while usually only the weights are trained in the RBN. The training formulation has been thoroughly discussed in [18].

In our implementation (Fig. 4), the mean-squared-error (MSE) over the four subfuzzy systems, with the output of each
standing for the activation level of one function, is used to
decide whether or not to stop training. The MSE is defined as

$$\text{MSE} = \sqrt{\frac{1}{N_t} \sum_{t=1}^{N_t} \sum_{c=1}^{N_c} (O_{t,c} - d_{t,c})^2}$$

where

- $N_c$: number of categories (= 4);
- $N_t$: number of trials (= 80);
- $O_{t,c}$: activation output level of trial $t$ of the $c$th category component, for example the case $c = 1$ corresponds to the function of supination, so $O_{t,1}$ is OUT_SP in Fig. 4 at trial $t$;
- $d_{t,c}$: is the desired output of $O_{t,c}$ (= 0.1 or 0.9).

Here $N_c$ is not included under the denominator because the word “mean” in MSE is indicating averaging the error across trials only and we have calculated the MSE in neural network counterpart in the same way. According to the tests performed on the system, the stopping MSE is usually set to 0.1–0.3. All four outputs are used in order to generate the “selected function” although the subfuzzy systems are independent before they are fed to the “output logic.” In practice the MSE values are not necessarily the same across outputs and hence an average or total MSE is used.

**B. Feature Selection**

Most of the features used in our fuzzy system are the same as
that of Hudgins et al. [5], except that SSC was not adopted as it
proved to be insignificant. In reproducing the ANN classification
scheme of Hudgins et al. [5], we found that classification performance was not improved and even deteriorated for some subjects with the involvement of the SSC. In this work, the EMG signal is divided into several adjacent segments. Each time segment duration is 40 ms and six segments are used, so that a total of $4 \times 6 = 24$ features are extracted from 240 ms of EMG signal. The features used are: MAV1, MAVSLP1, ZC1, WC1, MAV2, ..., WC2, MAV3, ..., WC3, MAV4, ..., WC4, MAV5, ..., WC5, MAV6, ..., WC6. In comparison, five segments (i.e., totally 200 ms) of EMG signal were used by Hudgins et al. [5]. To prevent any perceivable delay in real-time prosthetic control, the function action should be identified within 300 ms from the onset of the EMG; hence feature extraction may not be performed over any longer period. In this paper, we use six segments (i.e., 240 ms) instead of five to collect more information.

For comparison the reproduction of the ANN approach is also based on six segments of EMG.

**C. System Structure, Training, and Classification**

The schematic structure of the fuzzy approach is similar to
that of the ANN approach as in Fig. 1, except that the ANN block is replaced with the fuzzy logic system. Fig. 4 illustrates the fuzzy approach. The block “Fuzzy Logic System” is composed of four parallel subsystems similar to that in Fig. 3 generating OUT_EX, OUT_FL, OUT_PR, and OUT_SP, which represent the activation level of each function, respectively. The output logic generates the final classification result.

The “Basic Isodata for Initialization” block is activated in order to initialize the system’s fuzzy sets before the training phase. The desired outputs of OUT_EX, OUT_FL, OUT_PR, and OUT_SP are set to 0.9 for the selected function and 0.1 for others in the training phase. The four sub-systems are trained simultaneously with presentations of the whole training set. When MSE is lower than a preset threshold [usually 0.1–0.3, refer to Fig. 5(b)], the training stops.

After the system is trained, the initialization block is de-
activated, and the whole system is ready for classification. The “output logic” block finds the maximum value among OUT_EX, OUT_FL, OUT_PR, and OUT_SP. If this value is greater than 0.3, then the corresponding function is selected; otherwise, no function is activated. This threshold was opti-
mal determined with many experiments by the following optimality criteria: there is one but only one output among the four that exceeds the threshold. It should be noted that the OUT_EX, OUT_FL, OUT_PR, and OUT_SP have relative meaning rather than absolute, and they should not be regarded or interpreted as probabilities.

**D. Convergence Problem and Initialization of the Fuzzy System**

Generally speaking, the ANN is able to converge well due to the global nature of the sigmoid neuron-activation function used. However, with a Gaussian-shaped function no quick or good convergence is guaranteed without appropriate initialization.

It has been determined by Lippman [14] that a three-layer
ANN is able to have a decision region of convex-open or closed
regions. Selected carefully, good features tend to gather the
items belonging to the same class into one or more clusters in
the feature space, such that they tend to reflect their similarities
quantitatively. The basic Isodata algorithm [3] is able to cluster
items automatically. If cluster parameters are used to initialize
the fuzzy sets, each fuzzy set will be able to represent all or
part of the items in the same class at the very beginning, so
that blindness to initial definition of fuzzy sets is avoided and
convergence is accelerated in the next training stage. Based
on this idea, the fuzzy set’s centroid and width was initialized
using the results of the following Basic Isodata algorithm:

1) Randomly choose (24-dimensional) training data inputs
as initial values for the means $m_1, m_2, \ldots, m_c$, where $c$
is the desired cluster number, usually varying from 8 to
14 in this problem.
Fig. 4. Schematic diagram of fuzzy approach for EMG classification (MAV1 through WC6 are the 24 segmented features extracted; the initialization of fuzzy sets in fuzzy logic system is activated only once with clusters generated from Basic ISODATA algorithm; fuzzy logic system is trained and generates four outputs representing the possibility of each function; Output logic makes choice among these four outputs.)

Fig. 5. Mean squared error over training set and classification error rate over test set vs. training epoch number. (a) SIG-ANN system, training step size = 0.25; (b) ISO-FUZ system, training step size = 0.015. Isodata cluster number = 12, rule = 12 x 4. Subject #: no. 1.

Loop:
2) Classify all the \( \eta \) training samples by assigning them to the class of the closest mean.
3) Recompute the means through the averages of the samples in their class.
4) If any mean changed value, go to Loop; otherwise, go to step 5).
5) Assign the mean of each cluster to \( \bar{x}_i \), the centroid of the corresponding fuzzy set; compute average radii of each cluster and assign these values to \( \sigma_i \), the widths of the fuzzy sets.

Initialization as above proved to be a very effective way to ensure good convergence in the training process. Its strength is demonstrated in Fig. 5(b). In this example, the fuzzy system reached maximum classification performance level (lowest classification error rate) after only the fourth training epoch with the appropriate initialization.

To decrease the chance of cluster crossing over different classes, and to generate adequate rules in the fuzzy system, more than four clusters (four was the class total) were generated with several clusters assigned to the same class. However, the number of clusters has to be decided through experiments and experience. The larger the training set size, the more clusters that can be generated with little effect from noise-polluted samples. The cluster number was set to be 8–14 according to the class number (4 classes) and training set size (20 for each class) in this problem.

III. EXPERIMENTAL RESULTS

We denote the EMG Neural Network classification method with sigmoid activation function proposed by Hudgins et al. [5] as SIG-ANN; and the EMG fuzzy system classification method proposed in this paper as ISO-FUZ, which is initialized with the Basic Isodata algorithm and trained with the back-propagation algorithm.

Data sets from four subjects were used in the experiments; they are the same as those used by Hudgins et al. [5]. The training vs. test set size was 20 vs. 20 for each function for each subject so that at least 160 trials composed of four function categories were required in each subject’s data set. Both the SIG-ANN method and the ISO-FUZ method were applied to the data sets. The (correct) recognition rates of the two approaches were measured, and are compared in Fig. 6. The rates are close, with ISO-FUZ being slightly higher. As a matter of interest, ISO-FUZ was also tested on the mean of the time-segmented features (number of features reduced from 24 to 4); ISO-FUZ, in general, gave classification results comparable to SIG-ANN based on the same mean features.

The training speeds of ISO-FUZ and SIG-ANN are almost the same, e.g., 23 s for SIG-ANN and 25 s for ISO-FUZ with our software running on a Pentium PC (see Table II for complexity per epoch and Section IV-B for training epochs). The training speed depends on the computational complexity of the training procedure, which will be further discussed.

Fig. 7 demonstrates the outputs from ISO-FUZ compared with SIG-ANN outputs for each test trial of the same subject’s
data set. The categories are mutually exclusive, and as explained in Section II-C the selected function corresponds to the largest of the four outputs. In Fig. 7, trails 41 through 60 are pronation, and there are quite a few misclassifications among these 20 trials or samples. The features chosen for the classification algorithm have difficulty in distinguishing pronation, and these samples are often misclassified. The figure suggests that ISO-FUZ tends to give more consistent and hence more “stable” results than SIG-ANN does, since the former output variations over the same class are obviously smaller than the latter. To support this, we define \( \bar{\sigma} \), an average variation over the correctly-classified trials as the measure of output fluctuations (hence reflecting the consistency):

\[
\bar{\sigma} = \sqrt{\frac{\sum_{c=1}^{N_c} \sum_{t=1}^{N_t} r_t(O_{tc} - \overline{O_{tc}})^2}{\left( \sum_{t=1}^{N_t} r_t \cdot N_c \right)^2}},
\]

where

- \( N_c \) total number of categories (= 4);
- \( N_t \) total number of trials (= 80);
- \( r_t \) 1 if trial \( t \) is correctly classified or 0, otherwise;
- \( O_{tc} \) trial \( t \)'s activation output level of \( c \)th category component, for example, \( c = 1 \) corresponds to the function of supination, so \( O_{t1} \) is OUT_SP in Fig. 4 at trial no. \( t \);
- \( \overline{O_{tc}} \) average activation output level of \( c \)th category component over all the correctly classified trials in the same class of trial \( t \), calculated by

\[
\overline{O_{tc}} = \frac{\sum_{t_1, t_2 \in \text{Same Class}} r_{t_1} O_{t_1c}}{\sum_{t_1, t_2 \in \text{Same Class}} r_{t_1}}.
\]

In the example depicted in Fig. 7, the average variation \( \bar{\sigma} \) over subject no. 1 is 0.0647 for SIG-ANN, and 0.0495 for ISO-FUZ. The average variations of subject no. 1 through no. 4 are given in Table III, and indicate better output consistency for ISO-FUZ.

When the decision threshold in the output logic is increased, the error rate of classification also increases due to more rejected trials. On the other hand, if the threshold is arbitrarily decreased, to say less than 0.3, it will not make much sense since the lower threshold is too close to the inactivate level (= 0.1). We investigated the relation between error rate and decision threshold, as depicted in Fig. 8. From the figure, it is seen that the threshold of ISO-FUZ can be varied from 0.3 to 0.5 without increasing the error rate, while the range is only from 0.3 to 0.4 for SIG-ANN. The results presented in Fig. 8 are representative of the four subjects. They demonstrate ISO-FUZ’s wider threshold tolerance compared to SIG-ANN.

The overtraining problem was investigated by experiments, results of which are illustrated in Fig. 5. It was found that ISO-FUZ is less sensitive to overtraining when compared with SIG-ANN. In both methods, the mean squared error (MSE) over the training set decreased as the training epoch increased, but the classification rate over the test set did not decrease monotonically, with the rate increasing slightly after reaching its minimum point. We define \( B \), the increased amplitude of error rate, as the measure of over-training side effect:

\[
B = \max_{R_E(n) \leq \text{Minimum}} \{R_E(3n) - R_E(n)\}
\]

where \( R_E(n) \) is the test error rate after \( n \) training epochs. Greater \( B \) implies a more serious over-training effect. \( 3n \) could be replaced with \( 2n \). But with \( 2n \), ISO-FUZ would be too good because \( B = 0 \). In reality, ISO-FUZ still has a little over-training problem. In the example of Fig. 5, \( B \) was 0.025 for SIG-ANN, and 0.013 for ISO-FUZ, i.e., the over-training side effect of SIG-ANN is double that of ISO-FUZ. The results of \( B \) for all the four subject are tabulated in Table IV. It can be seen that ISO-FUZ involves less overtraining.

### IV. DISCUSSION

The ISO-FUZ method proposed in this paper was compared with SIG-ANN over the same data set and the same feature...
The ISO-FUZ method requires human involvement at the Basic Isodata initialization stage in order to get the minimum inter-class cross-over. Hence it is not automatic to the same extent as SIG-ANN. This drawback comes out of the trade-off between performance and automation.

B. Training Time Comparison

The algorithms were implemented on a Pentium PC in Turbo C. The computational complexities of SIG-ANN and ISO-FUZ at each training epoch are tabulated in Table II. The analysis shows that ISO-FUZ time cost per epoch is about 12.7 times as much as that of SIG-ANN to reach optimal classification result. ISO-FUZ, however, needs much fewer epochs to converge: ISO-FUZ 6 epochs vs. SIG-ANN 80 epochs. Table V gives the statistics of a sample training process over the data set of subject no. 1, in which the error rate of ISO-FUZ is lower than that of SIG-ANN, even when the MSE of the former is larger than that of the latter. Taking the example in Fig. 5, ISO-FUZ can be a good classifier even with large training MSE (as large as 0.5). Further, practical training of ISO-FUZ is potentially quicker than that of SIG-ANN.

C. Over-Training Situations

In the example of Fig. 5, the side-effect of over-training the SIG-ANN, measured by $B$, the error rate’s increasing amplitudes, is twice that of ISO-FUZ. Generally speaking, there is little over-training side effect for ISO-FUZ, while SIG-ANN can more easily suffer from over-training. When over-trained, SIG-ANN can yield such complex boundaries (overfitting) between classes that the generalization performance is undermined. Fuzzy systems approximate the ideal classifier (if it exists) by many patches (Kosko) [12]. It is the “local” nature of these patches that makes a fuzzy system robust and insensitive to over-training.

Overtraining of SIG-ANN can also result from a small training set size relative to the number of hidden neurons, or each class not being well represented in the training set. This is also true for ISO-FUZ, with the number of hidden neurons replaced with the number of rules.

V. CONCLUSION

This paper has proposed a new fuzzy approach for classification of EMG patterns. The training process and classification results of the proposed method are superior to those of a neural network-based approach, primarily in that the fuzzy system gives more consistent classification results and is insensitive to over-training. These features may be attributed to the structure of the ISO-FUZ system and the “supervised” selection of fuzzy set initialization data by the Basic Isodata algorithm. More favorably, it can adopt expert experience at the same time.
by additional fuzzy units in parallel. The fuzzy approach described in this paper can be universally applied to other classification problems as well.

The classification rates by ISO-FUZ are subject-dependent as in the SIG-ANN method evaluated, although with slight improvement. We are aiming for 95% with appropriate feature selection.

ACKNOWLEDGMENT

The authors would like to thank Dr. B. Hudgins and Mr. K. Englehart at the Institute of Biomedical Engineering, University of New Brunswick (UNB), Canada, for providing the EMG data set and for their valuable comments on this project.

REFERENCES